Pakistan J. Zool., vol. 50(2), pp 787-790, 2018. **DOI**: http://dx.doi.org/10.17582/journal.pjz/2018.50.2.sc7

Short Communication

Nucleotide Sequence of the Hypervariable Pre-S2 Surface Antigen Region of Pakistani HBV Isolates

Muhammad Rasool[1,*], Muhammad Zahid[1], Muhammad Ismail Khan[1], Qadeem Khan[2] and Khalid Khan[1]

[1]*Department of Zoology, Islamia College University, Peshawar*

[2]*Govt. Degree College, Dargai Fort, Dargai, Khyber Pakhtunkhwa*

ABSTRACT

This research work was conducted to find out nucleotide variations in the pre-S2 region of the two clinically diagnosed Surface Antigen of Hepatitis B Virus (HBsAg). Pre-S2 region nucleotide sequence of the sample MR02 was found completely homologous with the Pre-S2 region of already reported subtype "adr4". This sample showed 1.2 to 20% variation with the other reported sequences. Nucleotide sequence of the Pre-S2 region of the second sample MR03 does not show 100% homology with any of the reported sequences. It differs 2.4 to `18.8 % from the other reported sequences. Further it was established that MR02 and MR03 nucleotide sequences differ from one another at four positions, all substitutions. The first difference is observed at 87th base where there is "C" residue in MR02 DNA sequence and an "A" in MR03 DNA sequence. The 2nd, 3rd and 4th difference is at base number 116, 137 and 146 respectively. In all the cases "C" residue in MR02 sequence is replaced by "T" residue in MR03 nucleotide sequence. The first change at nucleotide level does not cause any change at amino acid level, while the 2nd, 3rd and 4th substitutions at nucleotide level results an amino acid change. The phylogenetic studies established that the Pakistanian HBV Isolates closely resemble the "adrq" strain belonging to "adr" subtype. The study showed that the pre-S2 region is hypervariable and two reported sequences show 100% homology and thus cause change in the structure of Hepatitis B Surface Antigen (HBsAg). This variability reduces the efficacy of vaccines prepared to a particular clone of HBV isolates.

Article Information
Received 19 July 2017
Revised 14 October 2017
Accepted 22 December 2017
Available online 23 March 2018

Authors' Contributions
MR did the lab work and wrote the article. MZ critically reviewed the article. MIK, QK and KK helped in tabulation of data.

Key words
HBV, adr, HBsAg, PCR, DNA, Nucleotide sequence.

Viral hepatitis, especially caused by infection with hepatitis B virus (HBV), is a major public health problem. HBV is a significant cause of post transfusion hepatitis and a major cause of chronic hepatitis and hepatoma in South Asia and Japan (Ngaira *et al*., 2016). In Pakistan studies have shown 10 to 30 % endemicity (Mahmood *et al*., 2016). Almost 62% of the liver cancers in Pakistan are due to hepatitis B virus infection (Shah and Shabier, 2002). In Karachi 15 to 25 % of the population is HBsAg positive (Mahmood *et al*., 2016). Hepatitis B is 100 times more infectious than AIDS and responsible for 1-2 million deaths per year worldwide (Zakhari, 2013).

HBV was first discovered in 1965 (Blumberg *et al*., 1965). It is the smallest known DNA virus about 42 nm in diameter. HBV consists of a core containing the viral genome (about 3200 base pairs) bound to the core protein and its own DNA polymerase (Zhou *et al*., 2017). HBsAg is a complex molecule and is encoded by the Pre-S1 region, Pre-S2 region and the S gene. Serologically HBsAg has one group specific antigen determinant "a" and two sets of mutually exclusive determinants "d" or "y" and "w" or "r" resulting in four major serotypes adw, adr, ayw and ayr (Asad *et al*., 2015). "w" is commoner that "r" in USA, but "r" is commonest in Thailand (Liu *et al*., 2016). HBsAg subtle have distinct geographical, epidemiological and anthropological settings (Li *et al*., 2017). Subtypes adr and adw were mainly found in South Asia and Japan. Subtype adr is prevalent in South China, adw and ayw in India and ayw in Mediterranean countries while adw is common in the Netherland and ayw in Portugal (Li *et al*., 2017). HBV can be classified by the homology of entire nucleotide sequence into at least five genotypes with intra-group differences of less than 5.6 % and inter-group differences of more than 8%. Nucleotide substitution, characteristic of the five genotypes, are distributed on HBV DNA so evenly that genotypes can be achieved by sequencing only a few hundred nucleotides (Salarnia *et al*., 2017).

Pre-S2 region was selected for this study due to its hypervariability, having point mutations (El-Mowafy *et al*., 2017). Amino acid sequence encoded by the Pre-S2 region is found highly immunogenic. Of the total mutations occurring in the HBV genome, 50% occurs in the hypervariable regions including the Pre-S2. The mutations in the nucleotide sequence can affect the virus in two ways: first by interfering with the function of regulatory sequences and second by the changing the amino acid sequence of the encoded viral protein, effecting the clinical course of HBV infection (El-Mowafy *et al*.,

* Corresponding author: imrasool72@gmail.com
0030-9923/2018/0002-0787 $ 9.00/0

2017). The Pre-S2 region is present immediately upstream the S region, encoding for the S domain of HBsAg (Liu *et al*., 2016). Pre-S2 region is 165 base pairs (coding 55 amino acids) containing three distinct epitopes (Vratnica *et al*., 2015). Evidence accumulated so far suggests that the Pre-S2 peptide plays an important role in immunity to HBV, which might be involved in the stimulation of antibodies formation against HBsAg (Chen *et al*., 2016). Function of the Pre-S2 encoded amino acids is to bind with the polyalbumin of the liver cells (Chen *et al*., 2016). The HBV-polyalbumin complex is taken up by endocytosis and in this way HBV enters the cytoplasm or nuclei of the liver cells. It is reported that response to the Pre-S2 domain is independent of the S domain and it can assist S domain non responder mice in producing antibodies through the Pre-S2 specific T cell function (Li *et al*., 2017).

Significant heterogeneity (greater than 10%) reduces the efficacy of vaccines and passive immunity reagents (Liu *et al*., 2016). This also allows the carriers of Asian HBV isolates to go undetected with the current serological detection kits because most of these vaccines and kits are developed with European and American HBV isolates.

Materials and methods

Blood samples were collected from clinically diagnosed HBsAg positive patients from Services Hospital and Jinah Hospital, Lahore, Pakistan. EDTA was added to a final concentration of 7.0 mM and plasma extracted by centrifugation.

Viral DNA was extracted from the plasma using the published method (Persing *et al*., 1993). The isolated viral DNA was subjected to 20 cycles of PCR (94°C for 2 min, 50°C or 55°C for 1 min, 72°C for 1 min) amplification in a total volume of 25 µl (master mix 20 µl, Taq DNA polymerase 2 µl and HBV DNA 3 µl). The PCR was performed in limiting dilutions format. Limiting dilutions were carried out according to the published procedure (Brown and Simmond, 1995). This scheme uses a statistical method to predict that when the HBV DNA dilution employed contains only a single HBV DNA molecule then only 1 out of 5 PCR reactions will turn out to positive. The first set of primers (HepF1 and HepR1), designed by our lab group at CEMB Lahore Pakistan, with an optimized annealing temperature of 50°C was termed as "External Primers". This amplifies a region of 1.1 Kb of the HBV genome from nucleotide number 2363 to 283. The other set of primers (B2833S and B170AS) with an optimized annealing temperature of 55°C was taken from a published source (Wait *et al*., 2016). This pair was termed as "Internal Primers" which amplifies a region of about 0.58 kb from nucleotide number 2814 to 185. The oligonucleotide primer B170AS was further used for the DNA sequencing (Fig. 1).

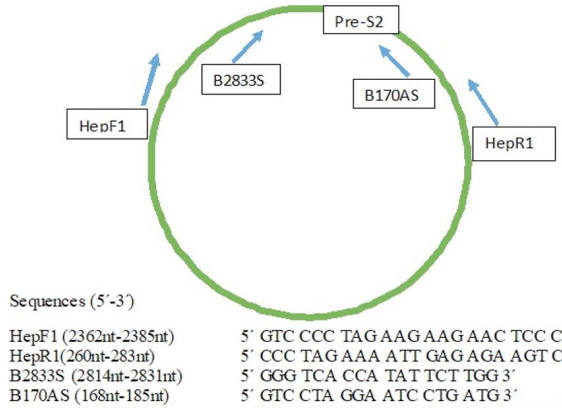

Fig. 1. Nucleotide sequence and diagrammatic sketch of primers used in PCR.

The PCR product of external primers was run on 0.8% low melt agarose gel and the bands of interest was purified from the gel slice through Wizard DNA purification system (Promega/Madison, WI, USA) or Sigma's Nucleiclean Kit. The gel purified DNA was cloned in the pGEM-T vector according to the manufacturer's instructions (Promega/Madison, WI, USA).

The gel purified PCR product of the external primers was sequenced through non-radioactive dideoxy chain termination method (Sanger *et al*., 1977). The sequencing reaction and staining of the gel was performed according to the manufacturer's instructions (Silver sequence kit; Promega/Madison, WI, USA).

Results and discussion

Blood samples were collected from two clinically diagnosed HBsAg positive patients (labelled as MR02 and MR03). According to the medical record of the hospital one patient (MR03) was in chronic condition while the other patient (MR02) was in the acute phase of the disease. The DNA isolated from the plasma of MR02 was amplified with the external primers and two out of five positive reactions were obtained at 1:4096 dilution (having 77 % probability that it has been amplified from a single HBV DNA molecule), DNA from one of the band was cloned in pGEM-T vector after gel purification. The insert was then removed from the vector through restriction digestion and sequenced using B170AS primer. The sequence is (5′ to 3′) 5′ATG CAG TGG AAC TCC ACA ACA TTC CAC CAA GCT CTG CTA GAC CCC AGA GTG AGG GGC CTA TAC TTT CCT GCT GGT GGC TCC AGT TCC GGA ACA GTA AAC CCT GTT CCG GCT ACT GCC TCA CCC ATA TCG TCA ATC TCC TCG AGG ACT GGG GAC CCT GCA CCG AAC3′. This sequence is completely homologous with the Pre-S2 region sequence of the HBV subtype adr4, while with the other reported sequences it shows a variation from 1.2 to 20 % (Fig. 2).

The second sample (MR03) DNA was subjected to PCR in dilution format and obtained one out of five positive (with external set of primers) at 1: 256 dilution. According to the scheme used for dilutions it (1/5 positive) has a probability of 90 % that it has been amplified from a single HBV DNA molecule. This band was gel purified and directly sequenced with B170AS primer.

The sequence is (5′ to 3′) 5′ATG CAG TGG AAC TCC ACA ACA TTC CAC CAA GCT CTG CTA GAC CCC AGA GTG AGG GGC CTA TAC TTT CCT GCT GGT GGC TCC AGT TCA GGA ACA GTA AAC CCT GTT CCG ACT ACT GTC TCA CCC ATA TCG TCA ATC TTC TCG AGG ATT GGG GAC CCT GCA CCG AAC3′. This sequence does not show 100 % homology with any of the reported sequences or with MR02. From MR02 this Pre-S2 region sequence of MR03 differs at four bases (Fig. 2). All of these four changes are substitutions. MR03 varies from the reported sequences by 2.4 to 18 %.

Both of these sequences were aligned with the reported sequences using computer software programme PCGENE CLUSTAL. According to the analysis both samples are in close resemblance with the Japanese clone adr4. The sample MR02 has 100 % homology with adr4 while MR03 has 96.7 % homology. MR02 homology with other subtypes follows the order adr>ayr>ayw>adw while MR03 shows the following order of homology with other known types adr4>ayr>ayw4>adw. A phylogenetic tree based on fourteen HBV Pre-S2 region sequences (containing MR02 and MR03) was prepared using computer software programme MEGA. The study of dendogram shows the possibility of Pakistanian HBV isolates to be evolved from adr ancestral subtype.

Purpose of the PCR amplification in limiting dilution format was to get homogenous product amplified from a single HBV DNA molecule. Heterogeneity in the HBV genome exists in the same subtype even prepared from a single donor (Ono *et al*., 1983). The most significant problem with the amplification of heterogeneous viral sequences by PCR is artefactual recombination. This generally occurs during the later cycles of PCR, where there is insufficient DNA polymerase to complete the synthesis of all primed DNA strands in the allotted time (Brown and Simmond, 1995). Partially extended DNA detaches from its template during the following denaturation step, then on cooling anneals to any available complementary strand, synthesis of partial strand is then completed by copying the wrong sequences. One out of five sequences are artificial recombinants if a template of about 300 base pair is subjected to 25 rounds of PCR amplification. The frequency of recombination increases with the increasing template fragment size (Meyerhans *et al*., 1990).

Significant heterogeneity has been observed in the sequence of the Pre-S2 region of HBV even within the same subtype (Vratnica *et al*., 2015). About 12 sequences of the Pre-S2 region of HBV have been reported so far and none of these sequences have 100 % homology with each other.

Fig. 2. Alignment of the Pre-S2 region sequences (two samples sequenced in the present study and twelve reported worldwide).

So it was perhaps not surprising to find that the Pre-S2 region sequence of Pakistanian HBV isolates described here were not 100 % homologous with each other. MR02 and MR03 DNA sequences reported in the present study vary from one another and also from the previously reported ones. MR02 and MR03 DNA sequences differ from one another by four bases. The first difference is observed at 87th base where there is "C" in MR02 and an "A" in MR03 sequence. The 2nd, 3rd and 4th variation is at base number 116, 137 and 146 respectively. In all these cases "C" of MR02 sequence is replaced by "T" in MR03 nucleotide sequence. It means that there is about 3 % variation between these two sequences. Due to these changes in nucleotide sequences, protein sequence of both the samples also vary from each other at three amino acid residues. Amino acid residue number 29 is serine in all of the reported HBV protein sequences and in the protein sequences of MR02 and MR03. There is a difference of only one nucleotide in MR02 and MR03 DNA sequence at this position. But this difference at DNA level does not affect translation as the amino acid remains the same. In the already reported sequences HBV Pre-S2 region protein sequences (12 sequences) there has been found a variation from alanine to valine at amino acid residue number 39. An alanine residue is present at the same position in the sequence of MR02 while a valine is present in the sequence of MR03 Pre-S2 region sequence. The amino acid residue number 46 also varies in both MR02 and MR03 due to one nucleotide substitution. In MR02 there is serine while in MR03 there is phenylalanine at the same position. Furthermore, an amino acid variation was observed at residue number 49. The published HBV protein sequences contain either threonine (8/12) or isoleucine (4/12) at this position. A threonine residue is present at the same position in MR02 while isoleucine in MR03. The effect of these changes on the structure or function of the protein is not clearly understood and needs further investigation.

Conclusion

The present data does suggest that most of the Pakistanian HBV isolates tend to belong to the adr subtype and appear to be derived from adrq ancestral strain belonging to adr subtype. Proper screening of the blood is a must before transfusion. Universal precautions when in contact with blood or blood contaminated objects should always be followed.

Acknowledgements

The authors are thankful to Naaz Abbas and Prof. Ziaur Rahman, Centre of Excellence in Molecular Biology, Punjab University, Lahore for their help and support.

Statement of conflict of interest

There is no conflict of interest among the authors.

References

Asad, A.M., Al-Ayed, M.S.Z., Aleraky, M. and Qurehsi, M.A., 2015. *Braz. J. Infect. Dis*., **19**: 525-528.

Blumberg, B.S., Alter, A.J. and Vinich, S.A., 1965. *J. Am. med. Assoc*., **191**: 541-546.

Brown, A.J.L. and Simmonds, P., 1995. *HIV Volume I: A practical approach* (ed. J. Karn). Oxford University Press, pp. 176.

Chen, J., Liu, Y., Zhao, J., Xu, Z., Chen, R., Si, L., Lu, S., Li, X., Wang, S., Zhang, K., Li, J., Han, J. and Xu, D., 2016. *PLoS One*, **11**: e0155654.

El-Mowafy, M., Elgaml, A., El-Mesery, M. and Elegezy, M., 2017. *Med. Virol*., **89**: 1559-1566.

Li, S., Wang, Z., Li, Y. and Ding, G., 2017. *Scient. Rep*., **7**: 1990.

Li, Y., Zhao, Y. and Wu, J., 2017. *PLoS One*, **12**: e0177713.

Liu, T., Shen, L., Yin, W., Wang, F., Zhang, G., Zheng, H., Dunzhu, D., Bi, S. and Cui, F., 2016. *Bing Du Xue Bao*, **32**: 156-160.

Mahmood, M.A., Khawar, S., Anjum, A.H., Ahmad, S.M., Rafiq, S. and Nazir, I., 2016. *Annls. King Edward med. Univ*., **10**: 18-24.

Meyerhans, A., Vartanian, J.P. and Wain-Hobson, S., 1990. *Nucl. Acids Res*., **18**: 1687-1688.

Ngaira, J.A., Kimotho, J., Miriqi, I., Osman, S., Nganga, Z., Lewembe, R. and Ochwoto, M., 2016. *Pan. Afr. med. J*., **24**: 315.

Ono, Y., Onda, H., Sasada, R., Igarashi, K., Sugino, Y. and Nishioka, K., 1983. *Nucl. Acids Res*., **11**: 1747-1757. https://doi.org/10.1093/nar/11.6.1747

Persing, D.H., Smith, T.F., Tenover, F.C. and White, J.T., 1993. *Diagnostic molecular biology: Principles and applications*. American Society for Microbiology, Washington, DC, pp. 375.

Salarnia, F., Besharat, S., Zhand, S., Javid, N., Khodabakhshi, B. and Moradi, A., 2017. *J. clin. Diag. Res*., **11**: 31-34.

Sanger, F., Nicklen, S. and Coulson, A.R., 1977. *Proc. natl. Acad. Sci. USA*, **74**: 5463-5467.

Shah, N.H. and Shabier, G., 2002. *J. Phys. Surg*., **12**: 386-371.

Vratnica, Z., Zehender, G., Ebranati, E., Sorrentino, C., Lai, A., Vujošević, D., Veo, C., Lo Presti, A., Cella, E., Giovanetti, M., Laušević, D., Galli, M., Ciccozzi, M. and Mugoša, B., 2015. *J. med. Virol*., **87**: 807-813.

Wait, S., Kell, E., Hamid, S., Muljono, D.H., Sollano, J. and Mohamed, R., 2016. *Lancet Gastroenterol. Hepatol*., **1**: 248-255.

Zakhari, S., 2013. *J. Gastroenterol. Hepatol*., **28(Suppl. 1)**: 18-25. https://doi.org/10.1111/jgh.12011

Zhou, H., Gewaily, D., Ahu, S.H., Preskill, C., Wang, Y., Zong, L., Zhang, J., Han, K.H., Wands, J., Li, J. and Tong, S., 2017. *Virus Res*., **235**: 86-95.